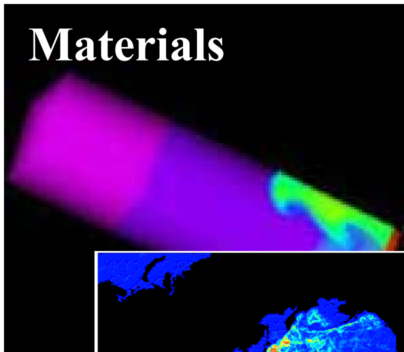
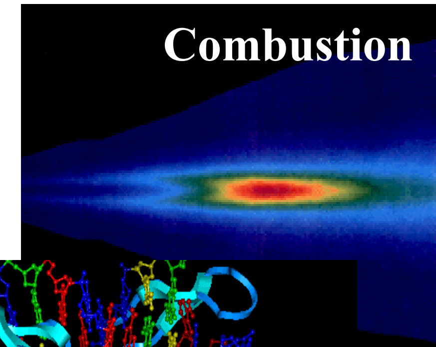
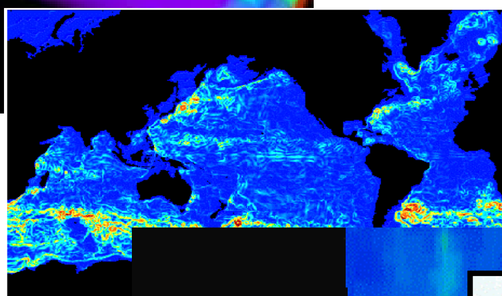
Scientific Computing

Critical to Discovery in Many Scientific Disciplines

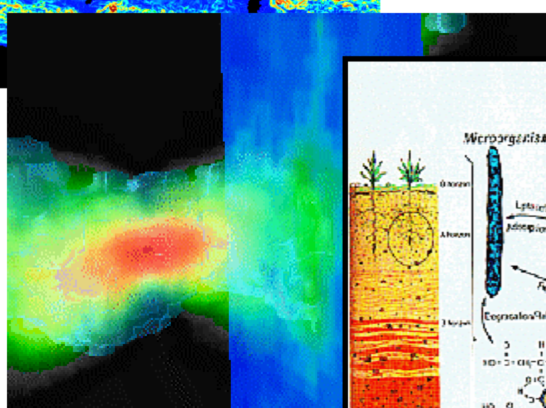




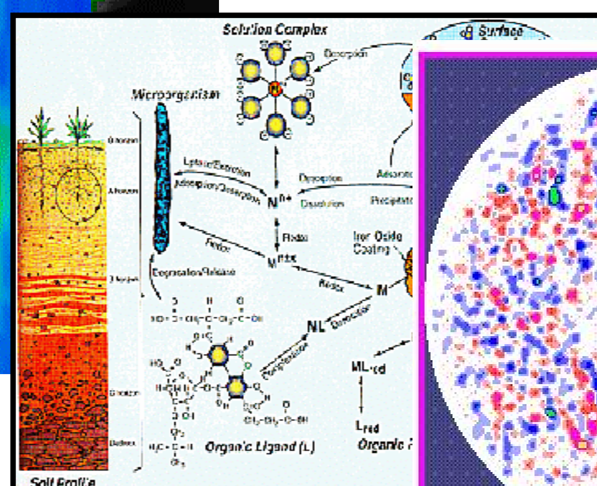
Many SC Programs Need Dramatic Advances in Simulation Capabilities To Meet Their Mission Goals



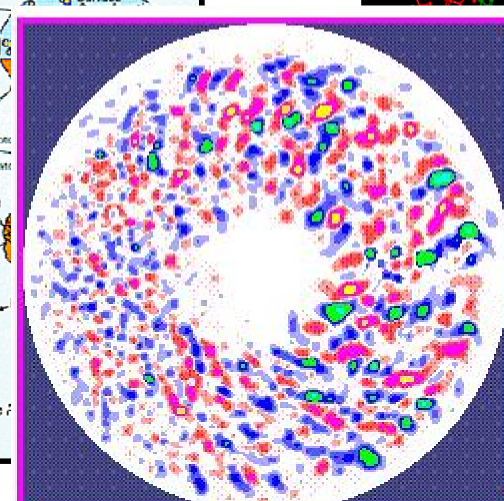
Global Climate



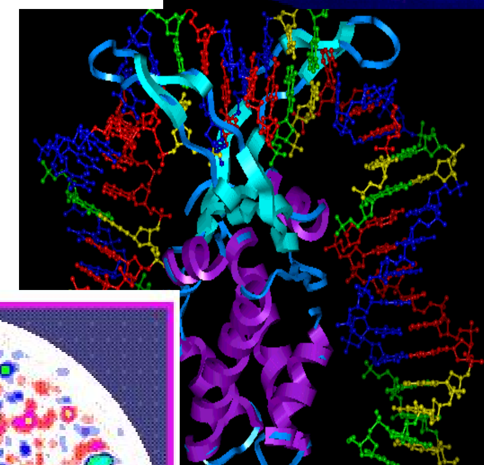
Components of Matter



Subsurface Transport



Fusion Energy



Health Effects, Bioremediation

Goal and Strategies

"Scientific Discovery through Advanced Computing"

Goal

- Promote scientific discovery throughout the Office of Science by exploiting advances in computing technologies

Strategies

- Create *Scientific Computing Software Infrastructure* that takes full advantage of terascale computing capabilities for scientific research
- Establish *Scientific Computing Hardware Infrastructure* that supports scientific research in the most efficient, effective manner possible
- Enhance collaboration and access to facilities and data through advances in networking technologies and development of electronic collaboratories

Programmatic Elements

"Scientific Discovery through Advanced Computing"

Hardware Infrastructure

Software Infrastructure

OPERATING SYSTEM

COLLABORATORIES DATA GRIDS

Data Analysis & Visualization

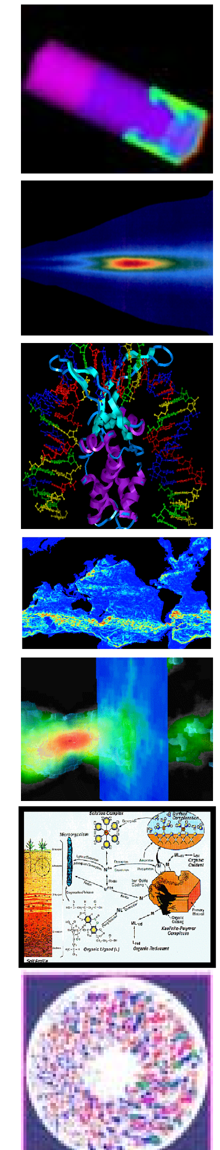
Mathematical Libraries

Programming Environments

Scientific Data Management

Problem-solving Environments

SCIENTIFIC SIMULATION CODES



BES, BER FES, HENP

ASCR

Software Challenges
Scientific Computing

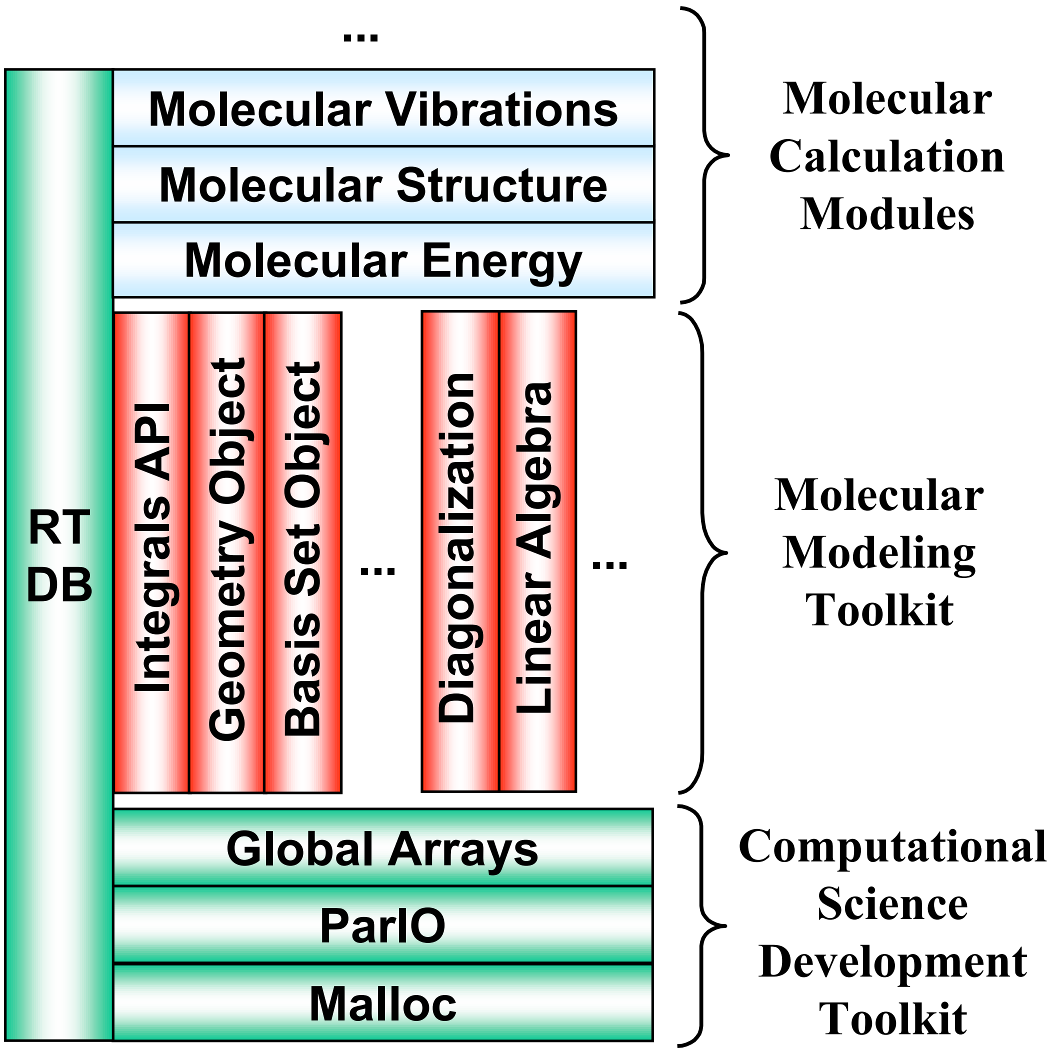
▲ Scientific Codes

- High fidelity mathematical models
- Efficient, robust computational methods and algorithms
- Well designed computational modeling and simulation codes
  - Readily incorporate new theoretical advances
  - Port from one computer to another with minimal changes

▲ Computing Systems and Mathematical Software

- Increased functionality in Vendor Operating Systems
- Computing systems software
  - Accelerate development and use of terascale scientific codes
  - Facilitate porting of software and codes among high-performance computers
  - Manage and analyze massive (petabyte) data sets, both locally and remotely
- Algorithms that scale to thousands-millions processors

An Example
Northwest Chem (NWChem)



NWChem

a major new modeling capability for molecular science

Molecular Electronic Structure

Molecular Dynamics

(Crystals)

> 600,000 lines of code and growing

Runs on ...

Cray T3D/E, IBM SP2, SGI SMP, NOWs, Sun and other workstations, X86 PCs (Linux)

Scales to ...

2,000+ processors

Developers ...

Core group (15) plus larger group (20) of world-wide collaborators

>100 person-years at PNNL alone

SC Software Infrastructure
A Major Software Challenge

Peak Performance is skyrocketing

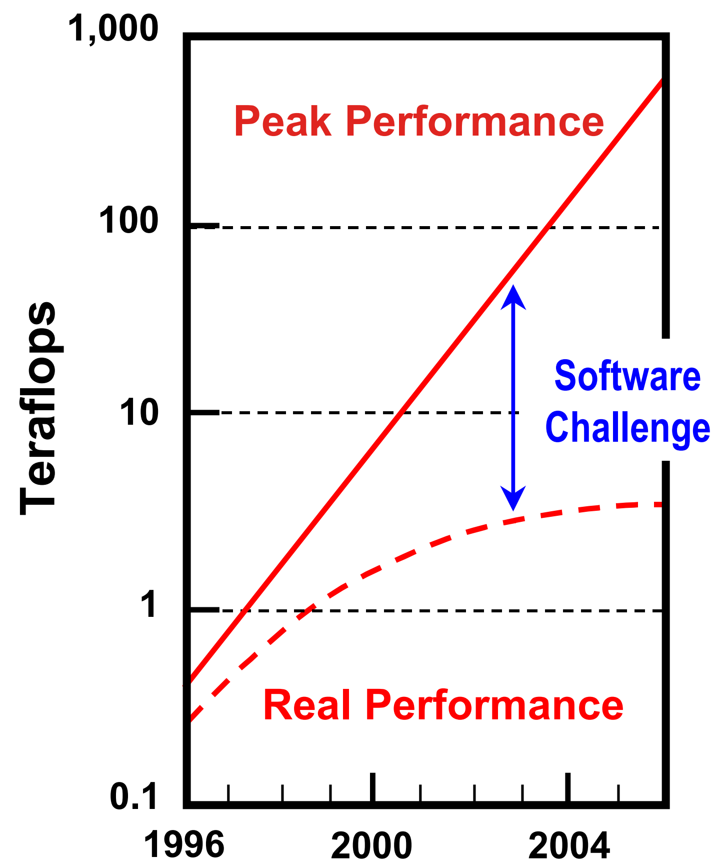
- In 90's, peak performance has increased 100x; in 00's, it will increase 1000x

But ...

- Efficiency has declined from 40-50% on the vector supercomputers of 1990s to as little as 5-10% on parallel supercomputers of today and may decrease further on future machines

Research challenge is software

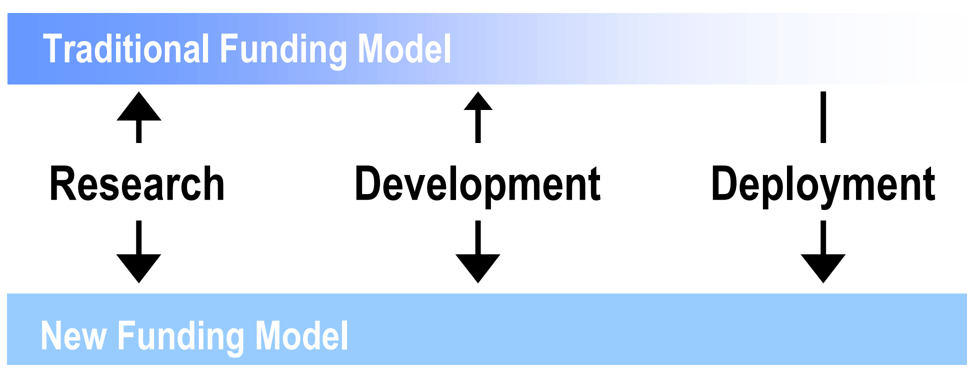
- Scientific codes to model and simulate physical processes and systems
- Computing and mathematics software to enable use of advanced computers for scientific applications
- Continuing challenge as computer architectures undergo fundamental changes



SC Software Infrastructure

Enabling Technology Centers

- **Teams of Mathematicians, Computer Scientists, Applications Scientists, and Software Engineers to ...**
  - Create mathematical and computing systems software to enable scientific simulation codes to take full advantage of the extraordinary capabilities of terascale computers
  - Work closely with Scientific Simulation Teams to ensure that the most critical computer science and applied mathematics issues are addressed in a timely fashion
  - Support the full software life cycle



SC Hardware Infrastructure
Robust, Agile, Effective & Efficient

- **Flagship Computing Facility**
  - To provide robust, high-end computing resources for *all* SC research programs
- **Topical Computing Facilities**
  - To provide the most effective and efficient computing resources for a set of scientific applications
  - To serve as a focal point for a scientific research community as it adapts to new computing technologies
- **Experimental Computing Facilities**
  - To assess new computing technologies for scientific applications

Why Topical Facilities?

Variation in Scientific Application Needs*

Code	Application	Time (TFLOPS-HRS)	Memory (TBYTES)	Storage (TBYTES)	Node I/O (MBYTES/S)
Cactus	Astrophysics	300	1.8	20	5
ARPS	Weather	25	0.25	16	18
MILC	Particle Physics	10,000	0.2	1	3
PPM	Turbulent Flow	500	0.5	54	6
PUPI	Liquids	150	0.1	0.2	3
ASPCG	Fluid Dynamics	5,000	0.5	50	3
ENZO	Galaxies	1,000	0.9	10	12
	Variation	**400x**	**18x**	**100x**	**6x**

* *From "High-level Application Resource Characterization," NSF/PACI National Computational Science Alliance, May 2000. Reported by permission of Dr. D. A. Reed, Director, National Center for Supercomputing Applications.*

Why Topical Facilities?

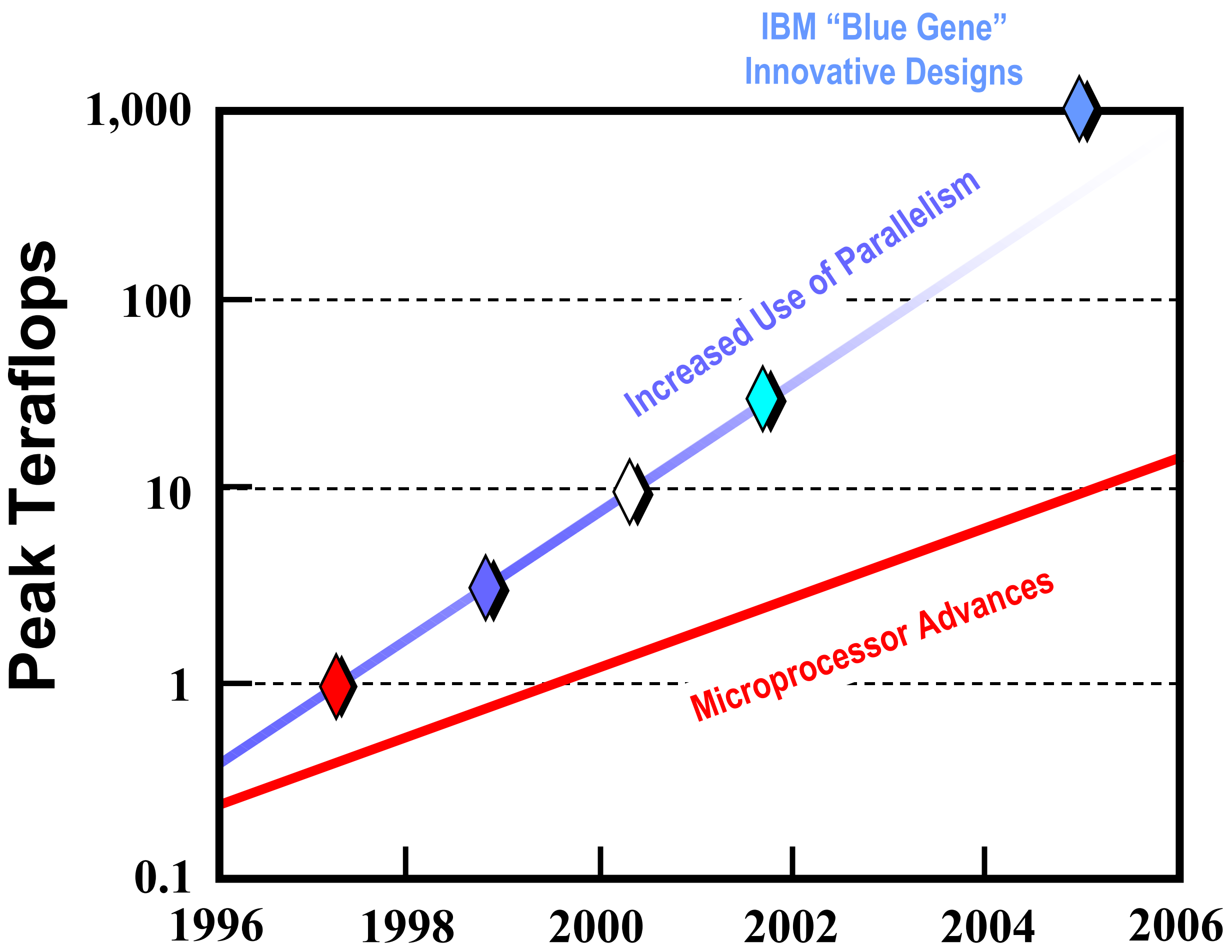
A Response to the Software Challenge

- **To provide the organizational framework needed for multidisciplinary activities**
  - Addressing software challenges require strong, long term collaborations among disciplinary computational scientists, computer scientists, and applied mathematicians

- **To provide the organizational framework needed for the development of community codes**
  - Implementation of many scientific codes requires a wide range of disciplinary expertise

- **Organizational needs will continue as computers advance to petaflops scale**

Dramatic Advances in Computing
Terascale Today, Petascale Tomorrow



MICROPROCESSORS

2x increase in microprocessor speeds every 18-24 months ("Moore's Law")

PARALLELISM

More and more processors being used on single problem

INNOVATIVE DESIGNS

Processors-in-Memory
HTMT

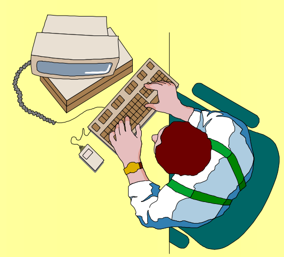
Elements of Topical Computing Facility

Computing Systems

Code Development Teams

Core Teams + Distributed Team Members

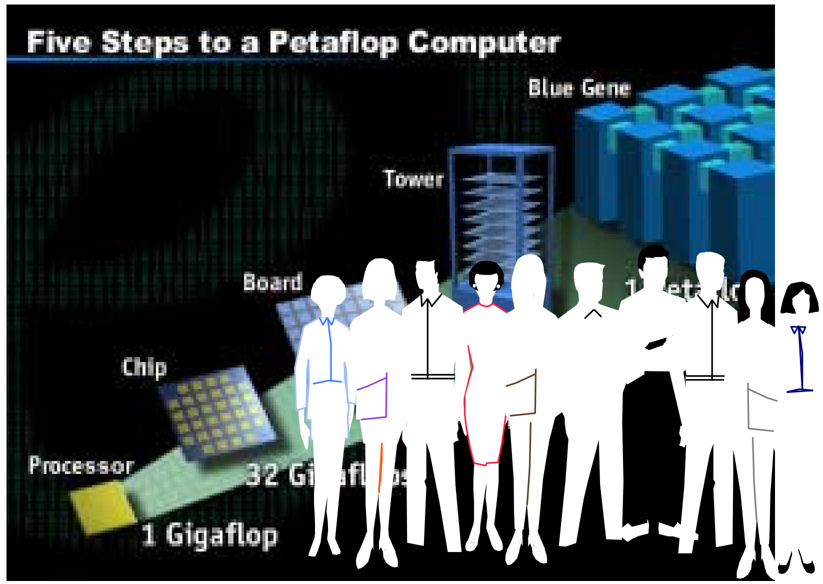
Scientific Application Users



Why Experimental Facilities?

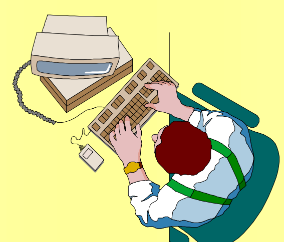
- **Need an organized approach for evaluation of new computing technologies (processors, switches, *etc*.)**
  - Although we are currently on a plateau vis-à-vis computer architectures, this will not last through the end of the decade
  - Examples of new approaches include PIM (processors-in-memory), HTMT (hybrid technology-multithread technology)
- **Need an organized approach for interacting with computer designers as early as possible**
  - Computer designers have many variables to consider – some beneficial for scientific computing, others not
  - Earlier the scientific community can provide input, the more likely the advice will be heeded

Elements of Experimental Computing Facility



Technologies Evaluation Teams

Core Teams + Distributed Team Members



Code Development Teams

@ Flagship Computing Facility

@ Topical Computing Facilities